Uses of Methods with Result Verification for Dealing with Uncertainty during MIMO Modeling and Simulation Process

Ekaterina Auer and Andreas Ahrens

University of Applied Sciences Wismar

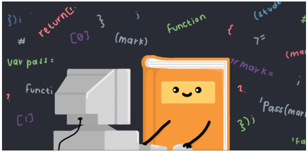
July 1, 2022

Guaranteed Results with Result Verification

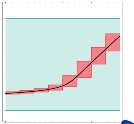
VERIFICATION — Are we building the product right? [Boehm]

FORMAL V
($\rightarrow$ model checking)

CODE V
($\rightarrow$ literate programming)



RESULT V
($\rightarrow$ interval arithmetic)



Principle: Prove mathematically the correctness of the computer result (fixed point theorems + set-based arithmetics)

Advantages: Account for rounding or conversion errors; propagate epistemic uncertainty

Disadvantages: Possibly too pessimistic ($\rightsquigarrow$ overestimation)

Approaches: Interval, affine, Taylor model, ... based methods

Result Verification: Applications

Computer assisted proofs

Smale's 14th: Do the properties of the Lorenz attractor exhibit that of a strange attractor?

Answer: Yes, proved by W. Tucker in 2002 with intervals

Other application areas

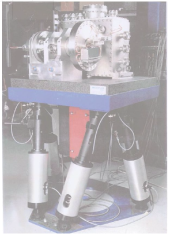
→ Computer graphics [Luther,Stolfi]
→ Finance/decision-making [Hu,Tsao]
→ Imprecise probability [Kreinovich,Ferson]

Main area: Engineering

→ Robotics [Jaulin,Merlet]
→ Chemical engineering [Stadtherr]
→ Particle accelerators [Makino,Berz]
→ Control theory [Walter,Rauh]
→ ... many more ...

And now: Wireless communications!

http://www.cs.utep.edu/interval-comp/



A positioner for the ESRF, Merlet

Focus: MIMO (**M**ultiple **I**nput **M**ultiple **O**utput) Systems

Improve the channel capacity/integrity without increasing the channel bandwidth or the transmit power

Method: Multiple data streams are transmitted on the same frequency band and at the same time

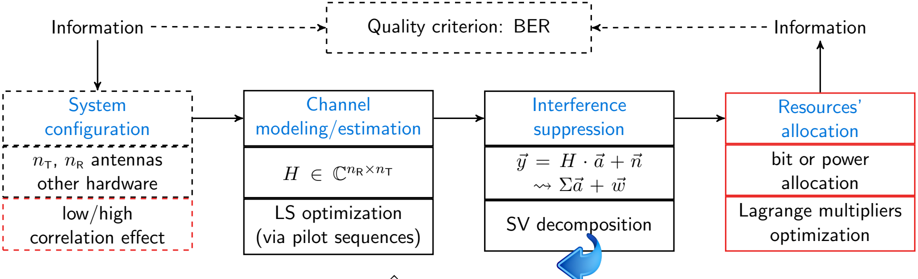
Separation: Spatial, for example, multiple antennas at the transmitter and receiver side at different locations

Correlation effect: Caused by the proximity of the multiple antennas; transmit-to-receive paths might become too similar!

Channel capacity: The information theoretic limit on the bit ratio (BER)

BER: The number of bits per second that can be transmitted through a physical channel error free

Modeling and Simulation of a MIMO Digital Channel



Interference suppression → $\hat{L}$ independent, weighted SISO links

(pre: $\vec{z} = V\vec{a}$, post: $\vec{u} := U^\dagger \vec{z} = U^\dagger \left(U \Sigma V^\dagger\right) V \vec{a} + U^\dagger \vec{n} = \Sigma \vec{a} + \vec{w}$)

$u_l = \lambda_l a_l + w_l$ for $l = 1 \ldots L$ (ideally non-interfering)

Resources allocation: L number of activated layers, $P_s^{(l)}$ transmit power, M_l constellation size

Each stage might be affected by uncertainty and numerical errors!

Uncertainty

MIMO achieve high capacity gains under perfect channel state information

Imperfect knowledge might be due to

Channel estimation stage:

→ channel estimation error at the receiver

→ limited feedback capability

Interference suppression/ Resources' allocation stages:

→ $\lambda_l = \sqrt{\xi_l}$ — errors in singular values of H

→ σ^2 — uncertain noise variance at the receiver side

Usually treated by traditional UQ techniques

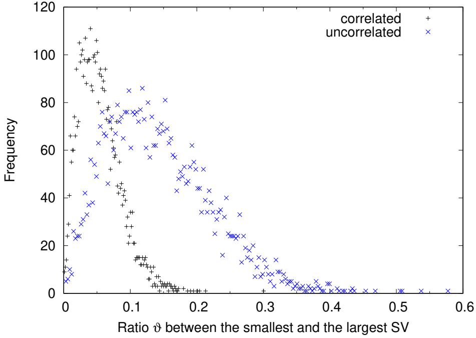
A combined treatment using verified techniques is possible

Good and Poor Scattering Conditions: Uncorrelated and Correlated Channel Realizations

Weights λ_l are not equal

Proximity might make this stronger! Indicator: $\vartheta = \dfrac{\text{the smallest } \lambda}{\text{the largest } \lambda}$

Illustration for a (4×4) MIMO channel (5000 realizations each):



Enclosing the BER for Uncertain Parameters

BER $P_b = \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \cdot \text{erfc}\left(\frac{\lambda_l}{2\sigma}\sqrt{\frac{3 \cdot P_s}{L(M_l - 1)}}\right)$

Task: Minimize the BER for uncertain $\lambda_l \in [\underline{\lambda}_l, \overline{\lambda}_l]$, $\sigma \in [\underline{\sigma}, \overline{\sigma}]$

$\rightarrow$ Minimize the upper bound!

Bound: $\frac{2}{\sum_{l=1}^{L} \log_2 M_l} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \cdot \text{erfc}\left(\frac{\underline{\lambda}_l}{2\overline{\sigma}}\sqrt{\frac{3P_s}{L(M_l - 1)}}\right)$

Minimize wrt. $P_s^{(l)}$ ($\rightsquigarrow$ power allocation) and L, M_l ($\rightsquigarrow$ bit allocation)

Power allocation: Largange multipliers + software with result verification

Bit allocation: Non-linear mixed-integer programming problem + software with result verification + power allocation

Power allocation: Problem Formulation

Idea: Assign more power to the layers with small weights! (L, M_l fixed)

$\rightarrow$ $P_s^{(l)} = \frac{P_s}{L}$ (equally distributed) $\rightsquigarrow \pi_l^2 \cdot P_s^{(l)}$ so that $\sum_{l=1}^{L} \pi_l^2 \cdot P_s^{(l)} = P_s$

Method: Constrained optimization with Lagrange multipliers

$$J(\pi_1 \ldots \pi_L, \mu) = \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \cdot \text{erfc}\left(\frac{\pi_l \lambda_l}{2\sigma} \sqrt{\frac{3 \cdot P_s}{L(M_l - 1)}}\right) + \mu \left(\overbrace{\sum_{l=1}^{L} \pi_l^2 - L}^{\text{constraint}} \right) \xrightarrow[\pi_l, \mu]{} \min$$

Power allocation: Verified Solution

Possibility 1 Mix analytical and numerical techniques

Stationary points: From the nonlinear algebraic system

$$\frac{\partial J(\pi_1 \ldots \pi_L, \mu)}{\partial \pi_l} = -\frac{2k_l}{\sqrt{\pi}} \left(c_l \lambda_l e^{-c_l^2 \lambda_l^2 \pi_l^2} \right) + 2\mu\pi_l = 0, \quad \sum_{l=1}^{L} \pi_l^2 - L = 0$$

$$\text{with } k_l = \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \cdot \left(1 - \frac{1}{\sqrt{M_l}}\right) > 0,\ c_l = \frac{1}{2\sigma}\sqrt{\frac{3 \cdot P_s}{L(M_l - 1)}} > 0$$

Solve using software with result verification $\rightsquigarrow$

C-XSC `www2.math.uni-wuppertal.de/wrswt/xsc/cxsc.html`

The (bordered) Hessian can be shown to be built in such a way that a stationary point is a local minimum!

Possibility 2 Use global optimization directly (e.g., in C-XSC)

Possibility 2 is usually more afflicted by overestimation!

Overestimation: A MIMO Link with Four Antennas, Four Active Layers ($L = 4$)

MIMO: Frequency flat, $n_\mathsf{T} = n_\mathsf{R} = 4$, $T = 8$ bit/s/Hz, $P_s = 1$ W

A data set with $\lambda_1 \approx 1.903$, $\lambda_2 \approx 0.624$, $\lambda_3 \approx 0.212$, $\lambda_4 \approx 0.0692$

Strong correlation: $\vartheta \approx 0.036$

Results for optimal π_i^2 at SNR of 10 dB ($\sigma \approx 0.2236$)

Possibility 2: $\pi_1^2 \in [0.5884, 0.5886]$, $\pi_2^2 \in [1.9511, 1.9513]$,
$\pi_1^3 \in [1.3002, 1.3005]$, $\pi_4^2 \in [0.15, 0.17]$

Possibility 1: $\pi_1^2 \in 0.588503196_1^9$, $\pi_2^2 \in 1.9511663_5^7$,
$\pi_1^3 \in 1.30033103_3^6$, $\pi_4^2 \in 0.159999408_8^9$

The Bordered Hessian

$$\begin{pmatrix} 0 & 2\pi_1 & \cdots & 2\pi_L \\ 2\pi_1 & 2\mu + \frac{4k_1c_1^3\lambda_1^3}{\sqrt{\pi}}\pi_1 e^{-c_1^2\lambda_1^2\pi_1^2} & \cdots & 0 \\ 2\pi_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 2\pi_L & 0 & \dots & 2\mu + \frac{4k_Lc_L^3\lambda_L^3}{\sqrt{\pi}}\pi_L e^{-c_L^2\lambda_L^2\pi_L^2} \end{pmatrix}$$

$\frac{\partial^2 J}{\partial\pi_l\partial\pi_m} = 0$ for $l \neq m$, , $\mu > 0$ from $\frac{\partial J(\pi_1 \dots \pi_L, \mu)}{\partial \pi_l} = 0 \rightsquigarrow \frac{\partial^2 J}{\partial \pi_l^2} > 0$

$$(l+1) \times (l+1): \quad \begin{vmatrix} 0 & a_1 & \cdots & a_l \\ a_1 & d_1 & \cdots & 0 \\ \vdots & & \ddots & \\ a_l & 0 & \cdots & d_l \end{vmatrix} = -\sum_{i=1}^{l}\left(a_i^2 \prod_{k=1, k\neq i}^{l} d_k\right) < 0$$

$\rightsquigarrow$ a local minimum in $(\pi_1 \dots \pi_L)$; unique solution $\rightsquigarrow$ globality

Power Allocation: A Special Case (1)

Special case: $L = 2$ and $M_1 = M_2 = M$ ($\lambda_1 > \lambda_2$ as usual)

The system: $k = \frac{1}{\log_2 M} \cdot \left(1 - \frac{1}{\sqrt{M}}\right)$, $c = \frac{1}{2\sigma}\sqrt{\frac{3 \cdot P_s}{L(M-1)}}$

$$\begin{aligned}
\frac{\partial J}{\partial \pi_1} &= \frac{2k}{\sqrt{\pi}}\left(-c\lambda_1 e^{-c^2\lambda_1^2\pi_1^2}\right) + 2\mu\pi_1 = 0 \\
\frac{\partial J}{\partial \pi_2} &= \frac{2k}{\sqrt{\pi}}\left(-c\lambda_2 e^{-c^2\lambda_2^2\pi_2^2}\right) + 2\mu\pi_2 = 0 \\
\frac{\partial J}{\partial \mu} &= \pi_1^2 + \pi_2^2 - 2 = 0
\end{aligned}$$

can be reduced to $\pi_1 = \frac{\lambda_1}{\lambda_2} e^{-c^2\left(\pi_1^2(\lambda_1^2+\lambda_2^2)-2\lambda_2^2\right)} \cdot \sqrt{2 - \pi_1^2}$

Not optimal: Choosing $\frac{\pi_1}{\pi_2} = \frac{\lambda_2}{\lambda_1}$ ($\pi_1 = \lambda_2\sqrt{\frac{2}{\lambda_1^2 + \lambda_2^2}}$)

Power Allocation: A Special Case (2)

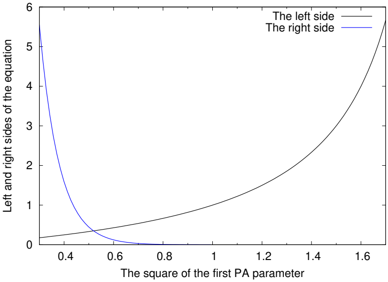
The intersection of a hyperbola and an exponential function

$$-1 + \frac{2}{2-x} = \left(\frac{\lambda_1}{\lambda_2}\right)^2 \kappa_1 e^{-\kappa_2 x}$$

with $\kappa_1 = e^{4c^2\lambda_2^2}$ and $\kappa_2 = 2c^2(\lambda_1^2 + \lambda_2^2)$.

Solution: $x \in (0, 2)$ (unique with e.g. Banach's theorem)

Example: $\lambda_1 = 4.341226$, $\lambda_2 = 2.178729$, $M = 16$, $\pi_1^2 \approx 0.51$



Bit allocation

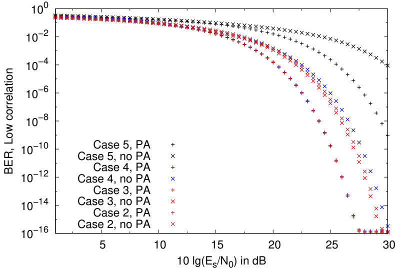
BER: $$P_b = \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \cdot \text{erfc}\left(\frac{\lambda_l}{2\sigma}\sqrt{\frac{3 \cdot P_s}{L(M_l - 1)}}\right)$$

→ M_l influences the overall BER

→ But also L (given throughput T)

case	L	layer 1	2	3	4
1	$L=1$	256	—	—	—
2	$L=2$	64	4	—	—
3	$L=2$	16	16	—	—
4	$L=3$	16	4	4	—
5	$L=4$	4	4	4	4

(throughput $T = 8$ bit/s/Hz)



Constraint: $\sum_{l=1}^{L} \log_2 M_l = T$

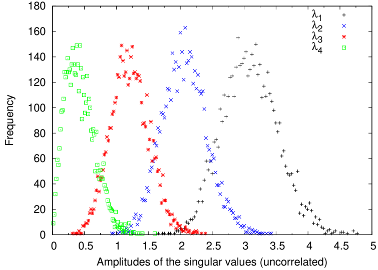
Solution: Brute force for small T and $\hat{L}$ possible (+ power allocation)!

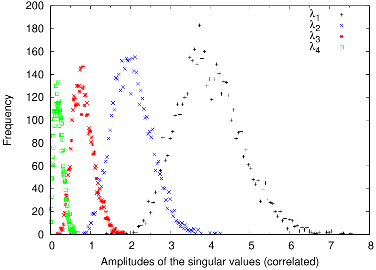
Example: A MIMO Link with Four Antennas

Simulation Settings

MIMO: Frequency flat, $n_\mathsf{T} = n_\mathsf{R} = 4$, $T = 8$ bit/s/Hz, $P_s = 1$ W

Two data sets with 5000 channel realizations each for correlated and uncorrelated case (simulated)



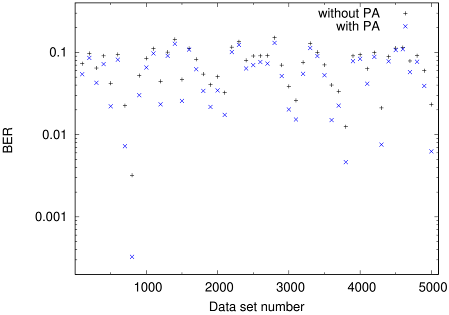


Results for the SNR 10 dB (corresponding to $\sigma \approx 0.2236$)

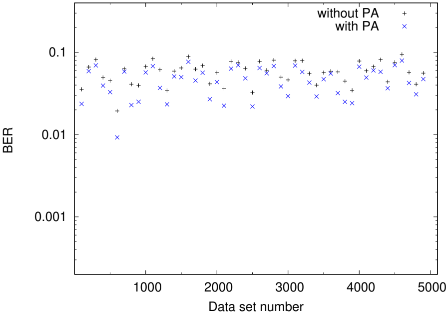
Power Allocation for Four Active Layers $(L = 4)$

Case $M_1 = 4,\ M_2 = 4,\ M_3 = 4,\ M_4 = 4$

Uncorrelated data set



Correlated data set



BER is reduced for each constellation of sigular values!

Bit Allocation for a (4×4) MIMO System

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \longrightarrow M_1 \geq M_2 \geq M_3 \geq M_4$$

([lowest, highest] BER for 5000 channel realizations)

	Layer M_1, M_2, M_3, M_4	BER (correlated)	BER-PA (correlated)	BER (uncorrelated)	BER-PA (uncorrelated)
			One active layer		
1	256, 0, 0, 0	[0.0023,0.1492]	the same	[0.0240, 0.13423]	the same
			Two active layers		
2	128, 2, 0, 0	[0.0022,0.1449]	[0.0001,0.1220]	[0.0232, 0.1304]	[0.0059, 0.1036]
3	64, 4, 0, 0	[$55\cdot 10^{-6}$,0.1103]	[$4\cdot 10^{-6}$,0.0959]	[0.0044, 0.0928]	[0.0007, 0.0749]
4	32, 8, 0, 0	[$10\cdot 10^{-7}$,0.0808]	[$2\cdot 10^{-7}$,0.0773]	[0.0002, 0.0599]	*[0.0001, 0.0556]*
5	16, 16, 0, 0	[$40\cdot 10^{-7}$,0.1092]	[$3\cdot 10^{-7}$,0.0981]	[$1.4\cdot 10^{-4}$, 0.06710]	[$1.1\cdot 10^{-4}$, 0.0589]
			Three active layers		
6	64, 2, 2, 0	[$8\cdot 10^{-4}$,0.1279]	[$8\cdot 10^{-6}$,0.1002]	[0.0127, 0.1121]	[0.0009, 0.0771]
7	32, 4, 2, 0	*[$12\cdot 10^{-6}$,0.0926]*	*[$2\cdot 10^{-6}$,0.0775]*	[0.0015, 0.0739]	[$6\cdot 10^{-5}$, 0.0533]
8	16, 8, 2, 0	[$11\cdot 10^{-6}$,0.1006]	[$5\cdot 10^{-6}$,0.0936]	*[0.0001, 0.06417]*	[$2\cdot 10^{-5}$, 0.0584]
9	16, 4, 4, 0	[$11\cdot 10^{-5}$,0.1015]	[$1\cdot 10^{-5}$,0.0972]	[$9\cdot 10^{-5}$, 0.0850]	[$1\cdot 10^{-5}$, 0.0785]
10	8, 8, 4, 0	[0.0001,0.1429]	[$7\cdot 10^{-5}$,0.1282]	[$2\cdot 10^{-5}$, 0.1048]	[$1\cdot 10^{-5}$, 0.0916]
			Four active layers		
11	32, 2, 2, 2	[0.0106,0.1532]	[0.0032,0.1255]	[0.0073, 0.1426]	[0.0005, 0.1129]
12	16, 4, 2, 2	[0.0071,0.1252]	[0.0023,0.1181]	[0.0006, 0.1099]	[$7\cdot 10^{-5}$, 0.1010]
13	8, 4, 4, 2	[0.0109,0.1665]	[0.0038,0.1529]	[$7\cdot 10^{-5}$, 0.1419]	[$4\cdot 10^{-5}$, 0.1344]
14	4, 4, 4, 4	[0.0414,0.2180]	[0.0228,0.2028]	[0.0014, 0.1909]	[0.0002, 0.1785]

All four layers should never be activated at the same time!

Interference Suppression – Another Approach

$$\boxed{\vec{y} = H \cdot \vec{a} + \vec{n}}$$

- $\rightarrow$ $\vec{y}$ received signal
- $\rightarrow$ $\vec{a}$ transmitted signal
- $\rightarrow$ $\vec{n}$ noise $\rightsquigarrow$ $\boxed{?}$
- $\rightarrow$ h_{ij} the fading coefficient between jth rec. / ith trans. antenna

Until now SVD: $\Sigma\vec{a} + \vec{w}$ with $\Sigma = diag\left(\sqrt{\xi_1}, \ldots, \sqrt{\xi_{\hat{L}}}\right)$

Pre-/Postprocessing: $\vec{z} = V\vec{a}$, $\vec{u} := U^\dagger \vec{z} = U^\dagger \left(U\Sigma V^\dagger\right) V\vec{a} + U^\dagger \vec{n} = \Sigma\vec{a} + \vec{w}$

$\rightarrow$ $u_l = \lambda_l a_l + w_l$, $\hat{L}$ independent SISO links with (unequal) $\lambda_l = \sqrt{\xi_l}$

Another possibility: GMD Decompose into $\hat{L}$ identical subchannels!

$\rightarrow$ $\hat{u}_l = \hat{\lambda}_l a_l + \hat{w}_l$ with $\hat{\lambda}_l = \sqrt[L]{\left(\prod_{i=1}^{L} \lambda_l\right)}$, $L = ?$

$\rightarrow$ Asymptotically optimal for high SNR (channel throughput, BER)

$\rightarrow$ Supposedly no trade-off between the capacity and BER

Conclusions

Results:

- → Problem solved by a mixed analytical/numerical technique with result verification
- → At least the weakest layer should be switched off
- → For correlated systems, resource allocation plays an especially important role
- → Best performance for two active layers

Future work:

- → Analyse the influence of the noise (σ)
- → Use GMD instead of SVD for obtaining equal weights — does the performance improve?

Thank you for your attention!